Formulae of Partial Reduction for Linear Systems of First Order Operator Equations

Branko Malešević [a], **Dragana Todorić** [b], **Ivana Jovović** [a], **Sonja Telebaković** [b]

[a] *Faculty of Electrical Engineering, University of Belgrade, Bulevar kralja Aleksandra 73, 11000 Belgrade, Serbia*
[b] *Faculty of Mathematics, University of Belgrade, Studentski trg 16, 11000 Belgrade, Serbia*

Abstract

This paper deals with reduction of non-homogeneous linear systems of first order operator equations with constant coefficients. An equivalent reduced system, consisting of higher order linear operator equations having only one variable and first order linear operator equations in two variables, is obtained by using the rational canonical form.

Key words: Linear system of first order operator equations with constant coefficients, n^{th} order linear operator equation with constant coefficients, sum of principal minors, the rational canonical form, the characteristic polynomial

1. Introduction

Let K be a field, V a vector space over K and let $A : V \longrightarrow V$ be a linear operator. **Linear system of first order A-operator equations with constant coefficients in unknowns x_i, $1 \leq i \leq n$,** is:

$$
\left\{
\begin{array}{l}
A(x_1) = b_{11}x_1 + b_{12}x_2 + \ldots + b_{1n}x_n + \varphi_1 \\
A(x_2) = b_{21}x_1 + b_{22}x_2 + \ldots + b_{2n}x_n + \varphi_2 \\
\qquad \vdots \\
A(x_n) = b_{n1}x_1 + b_{n2}x_2 + \ldots + b_{nn}x_n + \varphi_n
\end{array}
\right\}
\tag{1}
$$

for $b_{ij} \in K$ and $\varphi_i \in V$. We say that $B = [b_{ij}]_{i,j=1}^{n} \in K^{n \times n}$ is **the system matrix**, and $\vec{\varphi} = [\varphi_1 \ \ldots \ \varphi_n]^T \in V^{n \times 1}$ is **the free column**. If $\vec{\varphi} = \vec{0}$ system (1) is called homogeneous. Otherwise it is called non-homogeneous.

Let $\vec{x} = [x_1 \ \ldots \ x_n]^T$ be a column of unknowns and $\vec{A} : V^{n \times 1} \longrightarrow V^{n \times 1}$ be a vector operator defined componentwise $\vec{A}(\vec{x}) = [A(x_1) \ \ldots \ A(x_n)]^T$. Then system (1) can be written in the following vector form:

$$
\vec{A}(\vec{x}) = B\vec{x} + \vec{\varphi}.
\tag{2}
$$

A solution of vector equation (2) is any column $\vec{v} \in V^{n \times 1}$ which satisfies (2).

Powers of operator A are defined as usual $A^i = A^{i-1} \circ A$ assuming that $A^0 = I : V \longrightarrow V$ is the identity operator. By n^{th} **order linear A-operator equation with constant coefficients, in unknown x,** we mean:

$$A^n(x) + b_1 A^{n-1}(x) + \ldots + b_n I(x) = \varphi, \tag{3}$$

where $b_1, \ldots, b_n \in K$ are coefficients and $\varphi \in V$. If $\varphi = 0$ equation (3) is called homogeneous. Otherwise it is non-homogeneous. A solution of equation (3) is any vector $v \in V$ which satisfies (3).

2. Notation

Let $M \in K^{n \times n}$ be an n-square matrix. We denote by:

(δ_k) $\delta_k = \delta_k(M)$ the sum of its principal minors of order k $(1 \le k \le n)$,
(δ_k^i) $\delta_k^i = \delta_k^i(M)$ the sum of its principal minors of order k containing i^{th} column $(1 \le i, k \le n)$.

Let $v_1, \ldots, v_n$ be elements of K. We write $M^i(v_1, \ldots, v_n) \in K^{n \times n}$ for the matrix obtained by substituting $\vec{v} = [v_1 \ \ldots \ v_n]^T$ in place of ith column of M. Furthermore, it is convenient to use:

$$\delta_k^i(M; \vec{v}) = \delta_k^i(M; v_1, \ldots, v_n) = \delta_k^i(M^i(v_1, \ldots, v_n)).$$

The characteristic polynomial $\Delta_B(\lambda)$ of the matrix $B \in K^{n \times n}$ has the form

$$\Delta_B(\lambda) = \det(\lambda I - B) = \lambda^n + d_1 \lambda^{n-1} + \ldots + d_{n-1}\lambda + d_n,$$

where $d_k = (-1)^k \delta_k(B)$, $1 \le k \le n$; see [**2**, p. 78].

The companion matrix of a monic polynomial $p(\lambda) = \lambda^n + d_1 \lambda^{n-1} + \ldots + d_{n-1}\lambda + d_n$ is the matrix

$$C_p = \begin{bmatrix} 0 & 1 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & 1 \\ -d_n & -d_{n-1} & \ldots & -d_2 & -d_1 \end{bmatrix}.$$

The matrix C is in **the rational canonical form** if it is block diagonal, $C = \mathrm{diag}\big(C_{p_1}, C_{p_2}, \ldots, C_{p_k}\big)$, where C_{p_i} is the companion matrix of a monic polynomial $p_i(\lambda) = \lambda^{n_i} + d_{i,1}\lambda^{n_i - 1} + \ldots + d_{i,n_i - 1}\lambda + d_{i,n_i}$ of degree $n_i = \deg(p_i) \ge 1$, and $p_i \mid p_{i+1}$ for all i such that $1 \le i < k$.

3. Main Results

The reduction of system (2) to a **partially reduced system** will be divided into two steps.

Firstly, by using some basic properties of the rational canonical form, the initial system may be transformed into a proper system for further study.

Theorem 3.1 [*On rational canonical form* [**1**, p. 397]] *Any square matrix $B \in K^{n \times n}$ is similar to a unique matrix C in the rational canonical form. The matrix C is called the rational canonical form of the matrix B.*

Lemma 3.2 *Let C be the rational canonical form of the matrix B, i.e. there exists a regular matrix P such that $C = P^{-1}BP$. Then the system given in the form (2), $\vec{A}(\vec{x}) = B\vec{x} + \vec{\varphi}$, can be reduced to the system*

$$\vec{A}(\vec{y}) = C\vec{y} + \vec{\psi}, \tag{4}$$

where $\vec{\psi} = P^{-1}\vec{\varphi}$ is its free column and $\vec{y} = P^{-1}\vec{x}$ is a column of the unknowns.

The following lemma will be very useful in the second step of our reduction process.

Lemma 3.3 *If the matrix M has the form*:

$$M = \begin{bmatrix} b_1 & 1 & 0 & \dots & 0 \\ b_2 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ b_{n-1} & 0 & 0 & \dots & 1 \\ b_n & a_{n-1} & a_{n-2} & \dots & a_1 \end{bmatrix}$$

then it follows

$$\delta_k^1(M) = (-1)^k \left(\sum_{j=1}^{k-1} b_j a_{k-j} - b_k \right) \qquad (1 \le k \le n).$$

Proof. Observe that the function δ_k^1 is linear with respect to the first column of its matrix argument. The first column $[b_1 \ \dots \ b_n]^T$ of the matrix M is equal to $\sum_{j=1}^{n} b_j \vec{e}_j$, where $\vec{e}_j$ denotes j^{th} column of the identity matrix of order n. Therefore $\delta_k^1(M) = \sum_{j=1}^{n} b_j \delta_k^1(M; \vec{e}_j)$. It remains to prove that

$$\delta_k^1(M; \vec{e}_j) = \begin{cases} 0, & k < j; \\ -(-1)^k, & k = j; \\ (-1)^k a_{k-j}, & k > j. \end{cases}$$

Let $\alpha_{k,j}$ be an arbitrary nonzero minor in the sum $\delta_k^1(M; \vec{e}_j)$. Since it is a nonzero principal minor which contains 1 from the first column, it must include 1 from the first row. Hence, minor $\alpha_{k,j}$ must also include the element 1 from 2nd row and 3rd column. In the same manner we conclude that our minor contains 1 from i^{th} row and $(i+1)^{th}$ column for $i \in \{2, \dots, j-1\}$. Accordingly, minor $\alpha_{k,j}$ must contain the first j columns and rows.

For $k < j$ there is no such nonzero principal minor.

If $k = j$, then the only nonzero minor in the sum $\delta_k^1(M; \vec{e}_j)$ is the minor formed from the first j columns and rows. Expanding it by the first column, we easily conclude that its value is $(-1)^{k-1}$.

For $k > j$ the same reasoning implies that $\alpha_{k,j}$ must contain all columns and rows from 1st to j^{th}. Now, let $l+1$ be the smallest index greater than j such that $\alpha_{k,j}$ contains $(l+1)^{th}$ column. Consequently, the minor must include element a_{n-l}. We apply zig-zag argument again to deduce that our minor has a block formed from the last $n-l$ columns and rows. So, the only minor having the desired properties is the one obtained by deleting all columns and rows between $(j+1)^{th}$ and l^{th}. It follows that $k = j + n - l$ and $\alpha_{k,j} = (-1)^{j-1} \cdot (-1)^{n-l-1} a_{n-l} = (-1)^k a_{k-j}$. $\square$

Remark 3.4 The matrix M is the transpose of the doubly companion matrix introduced in [**4**], see also [**5-7**].

We will now prove the main theorem dealing with a special case when the rational canonical form of the system matrix has only one block.

Theorem 3.5 *Assume that the rational canonical form of the matrix B is exactly the companion matrix C of the characteristic polynomial $\Delta_B(\lambda) = \lambda^n + d_1 \lambda^{n-1} + \dots + d_{n-1}\lambda + d_n = \Delta_C(\lambda)$. Then linear system of first order A-operator equations* (1) *can be reduced to the partially reduced system*:

$$(\mathcal{R}_C)\begin{cases} \Delta_C(A)(y_1) = \displaystyle\sum_{k=1}^{n}(-1)^{k+1}\delta_k^1\big(C; A^{n-k}(\psi_1),\ldots,A^{n-k}(\psi_n)\big)\\[2mm] y_2 \quad = A(y_1) - \psi_1\\[1mm] y_3 \quad = A(y_2) - \psi_2\\[1mm] \quad\vdots\\[1mm] y_n \quad = A(y_{n-1}) - \psi_{n-1}\end{cases}, \tag{5}$$

where the columns $\vec{y} = [y_1 \ \ldots \ y_n]^T$ and $\vec{\psi} = [\psi_1 \ \ldots \ \psi_n]^T$ are determined by $\vec{y} = P^{-1}\vec{x}$ i $\vec{\psi} = P^{-1}\vec{\varphi}$ for a regular matrix P such that $C = P^{-1}BP$.

Proof. Applying Lemma 3.2, system (2) is equivalent to system (4), which can be presented in the following form:

$$\begin{cases} A(y_1) \ = y_2 + \psi_1\\[1mm] A(y_2) \ = y_3 + \psi_2\\[1mm] \quad\vdots\\[1mm] A(y_{n-1}) = y_n + \psi_{n-1}\\[1mm] A(y_n) \ = -d_n y_1 - d_{n-1}y_2 - \ldots - d_1 y_n + \psi_n\end{cases}. \tag{6}$$

We now consider the linear operator:

$$\Delta_C(A) = A^n + d_1 A^{n-1} + \ldots + d_{n-1}A + d_n I.$$

By eliminating $y_2,\ldots,y_n$ from the previous system, we obtain the operator equation:

$$\begin{aligned} \Delta_C(A)(y_1) = &\ \Big(A^{n-1}(\psi_1)\Big)\\ &+ \Big(A^{n-2}(\psi_2) + d_1 A^{n-2}(\psi_1)\Big)\\ &+ \Big(A^{n-3}(\psi_3) + d_1 A^{n-3}(\psi_2) + d_2 A^{n-3}(\psi_1)\Big)\\ &\ \vdots\\ &+ \Big(\psi_n + d_1\psi_{n-1} + d_2\psi_{n-2} + \ldots + d_{n-2}\psi_2 + d_{n-1}\psi_1\Big). \end{aligned} \tag{7}$$

Using formulae (6), (7) and Lemma 3.3 we complete the proof. $\square$

Remark 3.6 It is important to say that the reduction from the previous theorem is reversible, i.e. reduced system (5) implies system (1).

Remark 3.7 System (5) is called **partially reduced** because it has only one variable separated from the others by a linear n^{th} order A-operator equation.

We now continue with the general case of partial reduction.

Theorem 3.8 *Assume that the rational canonical form of the matrix B is the block diagonal matrix*

$$C = \mathrm{diag}\big(C_1, C_2, \ldots, C_k\big) \quad \text{for some } k, \ 2 \leq k \leq n,$$

where C_i are the companion matrices of the monic polynomials $\Delta_{C_i}(\lambda) = \lambda^{n_i} + d_{i,1}\lambda^{n_i-1} + \ldots + d_{i,n_i-1}\lambda + d_{i,n_i}$ of degree $n_i \geq 1$ and $\Delta_{C_i} \mid \Delta_{C_{i+1}}$ for all i, $1 \leq i < k$.

Let $\ell_1 = 0$ and let $\ell_i = \sum_{j=1}^{i-1} n_j$ for i, $2 \leq i \leq k$. Then linear system of first order A-operator equations (1) can be transformed into the conjunction of the partially reduced systems:

$$\bigwedge_{i=1}^{k} (\mathcal{R}_{C_i}), \tag{8}$$

4

where every subsystem $(\mathcal{R}_{C_i})$ *is of the form*:

$$
\left\{
\begin{aligned}
\Delta_{C_i}(A)(y_{\ell_i+1}) &= \sum_{k=1}^{n_i}(-1)^{k+1}\delta_k^1(C_i; A^{n_i-k}(\psi_{\ell_i+1}),\ldots,A^{n_i-k}(\psi_{\ell_i+n_i})) \\
y_{\ell_i+2} &= A(y_{\ell_i+1}) - \psi_{\ell_i+1} \\
y_{\ell_i+3} &= A(y_{\ell_i+2}) - \psi_{\ell_i+2} \\
&\ \ \vdots \\
y_{\ell_i+n_i} &= A(y_{\ell_i+n_i-1}) - \psi_{\ell_i+n_i-1}
\end{aligned}
\right\}.
$$

Columns $\vec{y} = [y_1 \ \ldots \ y_n]^T$ *and* $\vec{\psi} = [\psi_1 \ \ldots \ \psi_n]^T$ *are determined by* $\vec{y} = P^{-1}\vec{x}$ *and* $\vec{\psi} = P^{-1}\vec{\varphi}$ *for a regular matrix* P *such that* $C = P^{-1}BP$.

Proof. According to Lemma 3.2, system (2) is equivalent to the following system:

$$
\bigwedge_{i=1}^{k}\left(\vec{A}(\vec{y_i}) = C_i\vec{y_i} + \vec{\psi_i}\right),
\tag{9}
$$

where $\vec{y_i} = [y_{\ell_i+1} \ \ldots \ y_{\ell_i+n_i}]^T$ is a column of unknowns and $\vec{\psi_i} = [\psi_{\ell_i+1} \ \ldots \ \psi_{\ell_i+n_i}]^T$ is the free column of the linear system $\vec{A}(\vec{y_i}) = C_i\vec{y_i} + \vec{\psi_i}$. We now consider the linear operator:

$$
\Delta_{C_i}(A) = A^{n_i} + d_{i,1}A^{n_i-1} + \ldots + d_{i,n_i-1}A + d_{i,n_i}I.
$$

Therefore, by Theorem 3.5 every subsystem $\vec{A}(\vec{y_i}) = C_i\vec{y_i} + \vec{\psi_i}$ is equivalent to the partially reduced system $(\mathcal{R}_{C_i})$. We now obtain the equivalence between system (9) and the system $\bigwedge_{i=1}^{k}(\mathcal{R}_{C_i})$. $\square$

Remark 3.9 System $\bigwedge_{i=1}^{k}(\mathcal{R}_{C_i})$ is equivalent to the initial system (1). Every subsystem $(\mathcal{R}_{C_i})$ consists of a higher order linear A-operator equation, in only one unknown, and first order linear A-operator equations, in exactly two variables.

Remark 3.10 The initial system has a solution if and only if every higher order linear A-operator equation of the partially reduced system has a solution.

Remark 3.11 The above mentioned reduction formulae can be applied to two special types of non-homogeneous systems, to systems of differential and systems of difference equations with constant coefficients. The usual way of proving the reduction formulae for a non-homogeneous system of differential equations with constant coefficients is to use the Jordan canonical form. However, this method is somewhat complicated, see [**3,** pp. 166-178]. We give a better and more efficient alternative formulae of reduction for an arbitrary A-operator. Our method involves the rational canonical instead of Jordan form which makes reduction process much more easier to handle and enables us to find general solutions.

4. Examples

The reduction formulae in special cases of $n = 2$ and $n = 3$ will be given as examples.

Example 4.1 Consider the linear system of the first order A-operator equations, in unknowns x_1, x_2:

$$
\left\{
\begin{aligned}
A(x_1) &= b_{11}x_1 + b_{12}x_2 + \varphi_1 \\
A(x_2) &= b_{21}x_1 + b_{22}x_2 + \varphi_2
\end{aligned}
\right\}.
\tag{10}
$$

Depending on the block structure of the rational canonical form C of the matrix B, partially reduced system of system (10) has two possible forms.

$\mathbf{1^0}$ The rational canonical form C of the matrix B has one block. Then the partially reduced system has the form:

$$\left\{ \begin{array}{l} \Delta_B(A)(y_1) = \delta_1^1(C; A(\psi_1), A(\psi_2)) - \delta_2^1(C; \psi_1, \psi_2) \\ y_2 = A(y_1) - \psi_1 \end{array} \right\}.$$

$\mathbf{2^0}$ The rational canonical form C of the matrix B is diagonal matrix $C = \mathrm{diag}(C_1, C_1)$. Then system (10) is equivalent to the partially reduced system:

$$\left\{ \begin{array}{l} \Delta_{C_1}(A)(y_1) = \psi_1 \\ \Delta_{C_1}(A)(y_2) = \psi_2 \end{array} \right\}.$$

Example 4.2 We now deal with the system of first order A-operator equations in three unknowns x_1, x_2, x_3:

$$\left\{ \begin{array}{l} A(x_1) = b_{11}x_1 + b_{12}x_2 + b_{13}x_3 + \varphi_1 \\ A(x_2) = b_{21}x_1 + b_{22}x_2 + b_{23}x_3 + \varphi_2 \\ A(x_3) = b_{31}x_1 + b_{32}x_2 + b_{33}x_3 + \varphi_3 \end{array} \right\}. \tag{11}$$

Partially reduced system (8) of system (11) depends on the block structure of the rational canonical form C of the matrix B. The following cases are possible.

$\mathbf{1^0}$ The rational canonical form C of the matrix B has one block; then (8) has the form

$$\left\{ \begin{array}{l} \Delta_B(A)(y_1) = \displaystyle\sum_{k=1}^{3}(-1)^{k+1}\delta_k^1(C; A^{3-k}(\psi_1), A^{3-k}(\psi_2), A^{3-k}(\psi_3)) \\ y_2 = A(y_1) - \psi_1 \\ y_3 = A(y_2) - \psi_2 \end{array} \right\}.$$

$\mathbf{2^0}$ The rational canonical form C of the matrix B has two blocks $C = \mathrm{diag}(C_1, C_2)$, where C_1 and C_2 are matrices of order 1 and 2 respectively; then (8) has the form

$$\left\{ \begin{array}{l} \Delta_{C_1}(A)(y_1) = \psi_1 \\ \Delta_{C_2}(A)(y_2) = \delta_1^1(C_2; A(\psi_2), A(\psi_3)) - \delta_2^1(C_2; \psi_2, \psi_3) \\ y_3 = A(y_2) - \psi_2 \end{array} \right\}.$$

$\mathbf{3^0}$ The rational canonical form C of the matrix B has three blocks equal to C_1; then (8) has the form

$$\left\{ \begin{array}{l} \Delta_{C_1}(A)(y_1) = \psi_1 \\ \Delta_{C_1}(A)(y_2) = \psi_2 \\ \Delta_{C_1}(A)(y_3) = \psi_3 \end{array} \right\}.$$

REFERENCES

[1] F.M. Goodman, *Algebra: abstract and concrete*, Semisimple Press, Iowa City, 1998.

[2] Ф.Р. Гантмахер, *Теория матриц*, издание четвертое, Наука, Москва, 1988.

[3] И.Г. Петровский, *Лекции по теории обыкновенных дифференциальных уравнений*, издание седьмое, Издательство московского университета, Москва, 1984.

[4] J.C. Butcher, *A Mathematical Miniature*, Mathematical Miniatures and Apologies, NZMS Newsletter, No. 68, December 1996. (http://www.math.auckland.ac.nz/∼butcher/miniature/miniatureAndApology.html)

[5] J.C. Butcher, P. Chartier, *The effective order of singly implicit Runge Kuta methods*, Numerical Algorithms 20 (4) (1999) 269-284.

[6] J.C. Butcher, W.M. Wright, *Applications of doubly companion matrices*, Applied Numerical Mathematics 56 (3) (2006) 358-373.

[7] J.C. Butcher, *General linear methods for ordinary differential equations*, Mathematics and Computers in Simulation 79 (6) (2009) 1834-1845.